MEDINFO 2010
C. Safran et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-588-4-826

Analyzing effects of providing performance feedback at ward rounds on guideline adherence – The importance of feedback usage analysis and statistical control charts

Ameen Abu-Hanna[a], Saeid Eslami[a], Marcus J. Schultz[b], Evert de Jonge[c], Nicolette F. de Keizer[a]

[a]*Department of Medical Informatics, University of Amsterdam, Amsterdam, The Netherlands*
[b]*Department of Intensive Care, University of Amsterdam, Amsterdam, The Netherlands*
[c]*Department of Intensive Care, Leiden University Medical Center, Leiden, The Netherlands*

Abstract

Objective: Feedback to clinicians on their past performance is often aimed at increasing adherence to guidelines. We investigate how various analytical approaches influence the interpretation of adherence data. The analytical approaches vary in considering the actual or the intended use of the feedback, and whether outcomes are inspected over time. Material and Methods: At base line, a computerized decision support system was employed at the ICU bedside to increase adherence to a mechanical ventilation strategy. We intervened by providing feedback about adherence to the guideline at the daily ward rounds. The outcome measure was the percentage of ventilation time (V_T) in excess of the guideline's recommendation. Actual usage of the feedback was logged and data analysis was carried out using two approaches: classical statistics, and statistical process control (SPC) that inspect progress of an outcome over time. Design: Prospective, before/after study.
Results: The classical analysis stated that the percentage of ventilation time in excess of the guideline's recommendation decreased significantly due to the feedback (5% reduction, $p<0.001$). When SPC analysis of the outcome was applied, the effect was deemed not significant. When the actual delivery of feedback over time was also included it showed that the experiment does not allow for conclusive results. Conclusions: The concluded effect of providing feedback on adherence to a guideline depends on whether the actual usage pattern of the feedback and the inspection of the outcome over time are considered. Future evaluative studies should report on usage patterns and progression of outcomes over time.

Keywords:

Computerized Decision Support Systems, Provision of performance feedback, Adherence to Guidelines, Evaluative studies, Usage patterns, Statistical Process Control, Mechanical Ventilation, Ward rounds.

Introduction

Mechanical ventilation (MV) is one of the most important life supporting facilities in intensive care units (ICUs) [1]. Several complications, however, have been associated with the use of MV and various guidelines are used to limit the tidal volumes (V_T) applied to patients [2-4]. Such guidelines rely mainly on the predicted body weight (PBW) of the patient, which relies on the patient's height and gender.

Receiving too large V_T, however, is still the norm rather than the exception [5]. Explanations for this include the use of actual bodyweight instead of PBW, reluctance to accept potential hypercarbia and/or perceived increased sedation needs, and lack of knowledge [6-9]. In addition, it can be hypothesized that implementation of lower V_T is hampered by the fact that at times large V_Ts are simply left unrecognized.

Like in other settings, a clinical computerized decision support system (CDSS) can potentially contribute to increasing adherence to this guideline. We have hence implemented a CDSS at the bedside to alert physicians when the patient had a higher V_T than the guideline prescribes. In this paper we consider, aside from the bedside decision support, the effects of adding feedback on adherence at ward rounds, as they form a notable hospital context in which staff members work as a group [10]. In the sequel we refer to this ward round feedback as the "intervention", which itself can also be considered a form of decision support but provided in a new context. The objective of this paper is to contrast the results of three ways to analyze the data on he intervention: (A1) Using the traditional statistical approach, which is very common in the literature, (A2) applying the time-oriented approach of statistical process control (SPC), and (A3) Adding to A2 the patterns on actual usage of the feedback during the intervention period.

Materials and Methods

Study design

We performed a prospective before-after evaluation study on applied V_Ts to examine the impact of providing feedback at ward rounds on adherence (the intervention) to our local MV–protocol. The experiment was conducted during 2 consecutive phases in a 30–bed mixed medical–surgical ICU of a university hospital in the Netherlands.

Patient Data Management System

Since 2002, our ICU uses a commercial Patient Data Management System (PDMS, Metavision, iMDsoft, Sassenheim, The Netherlands). The PDMS is a point–of–care clinical information system, which runs on a Microsoft Windows platform and includes computerized order entry, automatic data collection, clinical documentation, electronic medication administration record, and a data storage repository. Mechanical ventilators are connected to the PDMS and all respiratory parameters are recorded every minute in the PDMS–database. Clinicians can see all fields related to artificial respiration, including applied V_Ts, by clicking on "respiratory tabs" in the PDMS.

Mechanical ventilation

The vast majority of patients are orotracheally intubated or do have a tracheotomy for MV; non–invasive MV is seldom applied. A written MV–protocol is available for all ICU–members, both on the intranet, and in printed form. In short, this protocol advises pressure–controlled (PC)–MV or pressure support (PS)–MV in all patients. There is a clear recommendation on V_T settings, stating that V_Ts have to be as low as 6 ml/kg PBW in all patients, irrespective of the presence of acute lung injury. With pressure–controlled (PC)–MV or pressure-support (PS)–MV, V_Ts are influenced by the applied airway pressures as well as the compliance of the patient's respiratory system. As compliance may change over time, healthcare workers need to continuously adjust the inspiratory pressure [11].

Recommended tidal volume

To calculate the recommended V_T (ml), the PBW (kg) is multiplied by 6. PBW was calculated by the following formula: in men, PBW = 50 + 0.91 * (height in centimeters – 152.4); in women, PBW = 45.5 + 0.91 * (height in centimeters – 152.4).

Ward round

The ward-round team consists of all ICU physicians and residents with possible additional specialists from other hospital wards like microbiologists or surgeons. The ward-round team has a meeting everyday to review patient progress. The PDMS is used to represent the patient conditions, complications and treatment plans. Patients are selected one by one and the related data are displayed by a beamer. The ward-round team discusses and reviews the patient charts, and decides upon the patients' treatment plan.

Intervention

At baseline, before the intervention, when an ICU–physician or ICU–nurse selected the "respiratory page" in the PDMS, the recorded V_T, maximum pressure (Pmax) and PEEP during the previous 60 minutes were queried. The system calculated the percentage of time that V_T was above the guideline recommended V_T. If the pressure support level was at its minimal value, the V_T was rendered "in range V_T". If V_T was above the guideline recommended V_T for more than 25% (15 minutes) of the previous 60 minutes, a pop–up window was shown displaying the guideline, patient's height, gender, PBW, as well as the percentage of time in which V_T was above 6 ml/kg PBW. In order not to disturb the users and to give them time to correct V_T, V_Ts were again queried and checked after at least two hours. The duration of this phase was almost 16 weeks.

In the intervention phase, in addition to showing the CDSS messages at the bedside, the daily ward round was used to display a list of the patients showing per patient the percentage of time in which V_T was above 6 ml/kg PBW. Patients with percentages below 25% were marked in green, percentages between 25-75% in yellow and above 75% in red. The duration of the intervention was also almost 16 weeks.

Usage of feedback

Although the feedback intervention was intended to be provided daily at the ward round, we logged its actual use: whenever the system displayed feedback at the ward round a record was created in the system to indicate this fact.

Outcome measures

The first outcome measure was the percentage of ventilation time in which the V_T was > 6 ml/kg PBW [4]. Mean V_T in excess of 6 ml/kg PBW over time was also measured. We considered all V_T measurements < 6 ml/kg PBW as if they were 6 ml/kg PBW. The third outcome was frequency of V_T measurements ≤ 6 ml/kg PBW. The unit of analysis is the V_T observation (not the patient).

Patients

This study included all ICU–patients who were mechanically ventilated for more than 24 hours in the ICU [5], were not on Adaptive Support Ventilation (which does not allow changes of V_T–settings by healthcare workers) and did not participate in other respiratory trials in which V_T was manipulated. V_Ts < 150 ml or >1500 ml were excluded as these were most likely measurement errors. When pressure support was at the lowest level, the measurement was considered correct (i.e., not in excess of 6 ml/kg PBW) regardless of its value.

Subgroups

We split up the whole ventilation time into spontaneous mode (pressure support ventilation [PSV]) and non–spontaneous mode (pressure control ventilation [PCV]).

Traditional statistical approach

Categorical variables in the before and after intervention groups were compared by X^2 testing, and continuous variables were compared by Student's *t* test or Mann-Whitney testing as appropriate. A *p*–value<0.05 was considered significant.

Statistical Process Control (SPC)

SPC and its primary tool – the control chart - is a branch of statistics that combines rigorous time series analysis methods with graphical data presentation, often yielding insights into the data more quickly and in a more understandable way than other statistical techniques [12-14]. Control charts can distinguish between common and special cause of variation. With common cause variation (noise), the variation is inherent in the process itself and the process is stable and predictable within certain limits. Special cause variation signifies that the

process is no longer stable or predictable and has changed (for better or worse). A control chart includes a plot of the data over time with three additional lines – the center line (usually the mean) and an upper and lower control limits, typically set at ±3 standard deviations (SD) from the mean. When the data points are, without any special pattern, within the control limits then the process is "in control" and stable. There are several rules that indicate when a special cause variation or special pattern has occurred on a control chart. We used the following four common rules [14]: one or more points above or below the control limit; a run of seven or more points on one side of the center line; two out of three consecutive points appearing beyond 2 SD on the same side of the center line; a run of seven or more points all trending up or down.

For analysis we used the X–MR chart (and not the attribute chart) due to the large size of observations per time point and the increased chance of false positive results [14-16]. Our chosen quality indicators (for guideline adherence) were calculated per two weeks, in order to allow for an adequate number of points, and plotted as points on the X–MR chart.

Results

Patients

During the study period 3,434,268 V_T–records (2,243,862 in the pre-intervention period vs. 1,190,406 in the intervention period) of 352 ventilated patients (202 vs. 150) were analyzed. Patient characteristics (age, height, severity of illness etc) before and after the intervention were similar.

Usage

In the first 5 weeks after the intervention the patients' list with the percentage of time in which V_T was above 6 ml/kg PBW was shown (35 times). After the fifth week, the CDSS was triggered only 1-2 times per week in the ward round (10 times in weeks 5-10, and 8 times in weeks 10-16 weeks).

Tidal volumes

Table 1 shows the outcome measures before and after the intervention. Using the traditional statistical approach the percentage of ventilation time with V_T in excess of 6 ml/kg PBW was shown to decrease significantly after intervening (5% reduction, $p < 0.001$). The average volume in excess of 6 ml/kg PBW remained the same after intervention.

Figure 1 shows the distribution of V_T (ml/kg PBW) by 0.25 ml/kg PBW in the two phases. The decrease in the percentage of time with $V_T \leq 6$ ml/kg PBW seems to come at the expense of the percentage of time with V_T between 6-8 ml/kg PBW, which decreased after the intervention.

Table 1 - Outcome measures reported for patients with >24 hours MV.

Mode	Intervention	Before	After	*p*
All	Mean of excessive V_T^{Φ}	1.3±1.7	1.3±1.7	ns**
	% of time > 6ml/kg PBW#	54	51	<.001*
Spontaneous (PSV)	Mean of excessive V_T	1.5±1.8	1.5±1.8	ns
	% of time > 6ml/kg PBW	58	54	<.001
Non-spontaneous (PCV)	Mean of excessive V_T	0.8±1.1	0.8±1.2	ns
	% of time > 6ml/kg PBW	43.5	44	ns

*chi square, **Student's *t* test, $^{\Phi}$Excessive V_T is tidal volume in excess of 6 ml/kg PBW over the whole phase. #To calculate this percentage we tolerated a discrepancy of 10% above the protocol's suggestion (i.e. 6.6 ml/kg PBW) before considering a measurement as being in excess of 6 ml/kg PB

Tidal volumes in (non)spontaneous ventilation mode

Table 1 also shows the outcome measures separately for spontaneous and non-spontaneous ventilation time. Surprisingly the overall effect of the intervention on spontaneous ventilation mode was larger than in the non-spontaneous mode. Results indicate that showing the patients' V_T information in the ward round did not change the percentage of ventilation time in excess of 6 ml/kg PBW during the non-spontaneous ventilation time. On the other hand, during the spontaneous ventilation time, the percentage of ventilation time with V_T in excess of 6 ml/kg PBW decreased after intervention (7% reduction, $p < 0.001$).

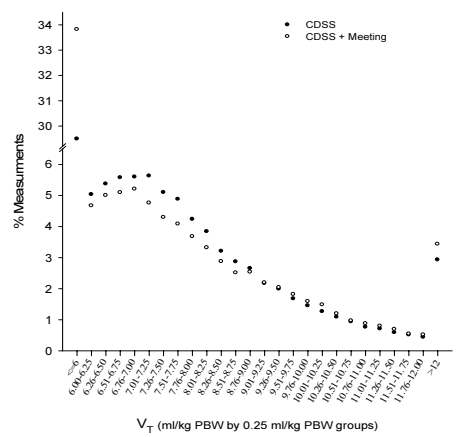

Figure 1 - V_T distribution (PSV mode).

Results based on statistical process control

Figure 2 shows control charts of the main outcome measures, also separately for spontaneous and non-spontaneous ventilation. In contrast to the traditional statistical methods, the control chart showed that the process did not change significantly after the intervention. Also in contrast to classical tests, the reduction pattern after intervention was only shown in the non-spontaneous mode.

Incorporating data on feedback usage

The control charts showed that the percentage of ventilation time with V_T in excess of 6 PBW did decrease steadily during the first 5 weeks but again increased and eventually became stable (The letter "A" in Figure 2 indicates the subgroup of measurements collected in week 5 and week 6). Our data on actual usage of feedback revealed that in these first 5 weeks the feedback delivery was provided daily, as intended. The increase afterwards in the outcome (indicating decrease in adherence to the guideline) is paralleled by the diminished use of feedback during the intervention.

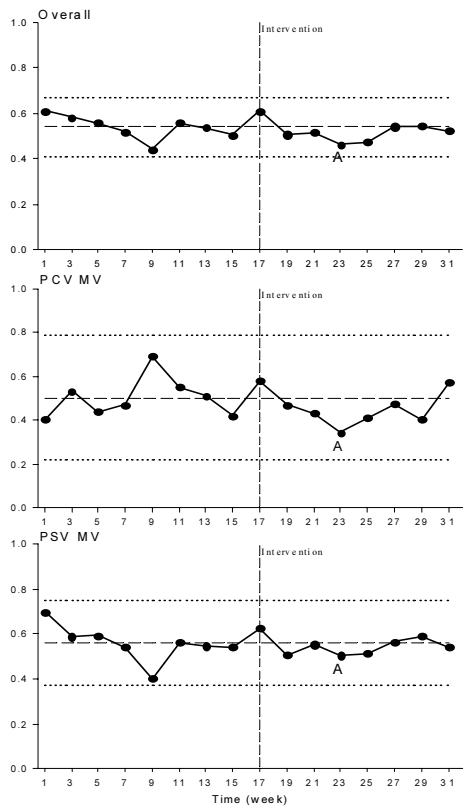

Figure - 2 Control charts of percentage of time > 6ml/kg PBW (overall, during spontaneous (PSV) and non-spontaneous (PCV) mechanical ventilation)

Discussion and Conclusion

Interpreting data pertaining to the effects of providing feedback concerning the performance of clinicians on adherence to a guideline provided three different pictures. In the traditional approach one would have concluded that the effect of our intervention was significant. If one takes into account the progression in the outcome measure over time but disregards the patterns of actual feedback delivery (and hence use) then the conclusion is that the intervention was not effective. If one in addition considers the usage pattern of the feedback then a nuanced picture emerges: as long as the feedback was provided daily, as planned, a decreasing pattern in the main outcome was observable. Once the frequency of feedback delivery has dropped the outcome started increasing again.

To understand the results it is important to point out the ambiguity of the term "intervention". There is the intended intervention (providing feedback daily) and the actual intervention (as was actually provided). Based on our experiments, it would seem too hasty to conclude that there is sufficient evidence (in terms of statistical significance) that our 16 weeks *actual intervention* was effective. An inspection of the SPC charts reveals why it arrived at a different conclusion than classical statistical tests: an apparent initial increase is counterbalanced by an increase in the sequel. Being sensitive to temporal progression, SPC is reluctant to declare statistical significance of this actual intervention (with reduced intensity of the feedback over the weeks). We believe that the SPC's prudent interpretation should prevail. By the same token concluding that the *intended intervention* is not effective is also unfounded. This is because this intended intervention was not implemented. The correct interpretation of the results is that the *intended intervention* does not enjoy enough quantitative support yet (as the usage pattern shows it was not implemented). However, inspecting the patterns of decrease in the first 5 weeks and "bouncing back" afterwards does provide qualitative evidence to the possible effectiveness of this intervention, which should be more properly implemented.

The same logic applies to subgroup analysis. While the control charts showed that the percentage of time > 6ml/kg PBW in the spontaneous mode was stable after the intervention, the traditional approach declared statistical significance.

There are two lessons (serendipitously) learned from our case-study for medical informaticians evaluating effects of IT interventions, such as decision support, on some (quality) indicator. First, measuring the progression over time of an indicator may suggest a different (and better) interpretation of the results than when time is not taken into account. The advantages of this interrupted time-series design is described in [17]. Second, measuring actual usage of an intervention may influence the interpretation of the results. These lessons are important because most of the medical informatics literature reports on results using the traditional approach. The few studies that did apply SPC analysis usually do not consider the possible difference between intended and actual use of the intervention.

Our results stress the importance of reporting or at least reflecting on the possible influence of a time-oriented approach and the CDSS usage on the results. After all, our study could have been reported in three incompatible ways.

References

[1] Pingleton SK. Complication of acute respiratory failure. Am Rev Respir Dis 1988; 137:1463-93.

[2] Eichacker PQ, Gerstenberger EP, Banks SM, et al. Meta-analysis of ALI and ARDS trials testing low tidal volumes. Am J Respir Crit Care Med 2002; 28:28.

[3] Richard J-D, Dreyfuss D, Saumon G: Ventilator-induced lung injury. Eur Respir J 2003; Suppl 22:42, 2s-9s.

[4] The Acute Respiratory Distress Syndrome Network: Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. N Eng J Med 2000; 342:1301-1308.

[5] Eslami S, de Keizer NF, Abu-Hanna A, et al. Effect of a clinical decision support system on adherence to a lower tidal volume mechanical ventilation strategy. J Crit Care 2009; doi:10.1016/j.jcrc.2008.11.006.

[6] Wolthuis EK, Korevaar JC, Spronk P, et al. Feedback and education improve physician compliance in use of lung-protective mechanical ventilation. Intensive Care Med 2005; 31:540-546.

[7] Rubenfeld G, Caldwell E, Hudson L. Publication of study results does not increase use of lung protective ventilation in patient with acute lung injury. Am J Respir Crit Care Med 2001; 163:A295.

[8] Young MP, Manning HL, Wilson DL, et al. Ventilation of patient with acute lung injury and acute respiratory distress syndrome: Has new evidence changed clinical practice? Crit Care Med 2004; 32:1260-1265.

[9] Kalhan R, Mikkelsen M, Dedhiya P, et al. Underuse of lung protective ventilation: Analysis of potential factors to explain physician behavior. Crit Care Med 2006; 34(2):300-306.

[10] Morrison1 C, Jones M, Blackwell A, Vuylsteke A. Electronic patient record use during ward rounds: a qualitative study of interaction between medical staff. Critical Care 2008; 12:R148.

[11] Campbell RS, Davis BR. Pressure-controlled versus volume-controlled ventilation: does it matter? Respir Care 2002; 47(4):416-24.

[12] Carey RG. How do you know that your chart is improving? Part I: Basic concepts in statistical thinking. J Ambulatory Care Manage. 2002;25(1):80-87.

[13] Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and healthcare improvement. Qual Saf Health Care. Dec 2003;12(6):458-464.

[14] Mohammed MA, Worthington P, Woodall WH. Plotting basic control charts: tutorial notes for healthcare practitioners. Qual Saf Health Care. Apr 2008;17(2):137-145.

[15] Carey RG. Improving healthcare with control chart: Basic and advanced SPC methods and case studies. Milwaukee: ASQ quality press; 2003.

[16] Wheeler DJ. Advanced topics in statistical process control. Knoxville, TN: SPC Press; 1995.

[17] Harris AD, McGregor JC, Perencevich EN, et al. The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics. J Am Med Inform Assoc. 2006; 13:16-23.

Address for correspondence

Ameen Abu-Hanna
Department of Medical Informatics
Academic Medical Center
University of Amsterdam
Meibergdreef 15
1105 AZ Amsterdam
The Netherlands.
a.abu-hanna@amc.uva.nl